View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Explore Bristol Research

OPEN ACCESS

University of BRISTOL

Taylor, N. (2017). Realised variance forecasting under Box-Cox transformations. *International Journal of Forecasting*, *33*(4), 770-785. https://doi.org/10.1016/j.ijforecast.2017.04.001

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
10.1016/j.ijforecast.2017.04.001

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at http://www.sciencedirect.com/science/article/pii/S0169207017300365. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms

# Realised variance forecasting under Box-Cox transformations

**Abstract**

The benefits associated with modeling Box-Cox transformed realised variance data are assessed. In particular, the quality of realised variance forecasts with and without this transformation applied are examined in an out-of-sample forecasting competition. Using various realised variance measures, data transformations, volatility models and assessment methods, and controlling for data mining issues, the results indicate that data transformations can be economically and statistically significant. Moreover, the quartic root transformation appears to be the most effective in this regard. The conditions under which the effectiveness of using transformed data are identified.

*Keywords:* Realised variance, Box-Cox transformation, forecasting competitions, loss function, reality check.
*JEL:* C22, C53, C58, G17.

## 1. Introduction

Power transformations, and more generally Box-Cox (BC) transformations, have long been recognised as an effective way of achieving well specified models with symmetric errors and stable error variance; see Tukey (1957) and Box and Cox (1964). More recently, attention has focused on assessing the out-of-sample performance of time-series models applied to BC transformed data. For instance, Bårdsen and Lütkepohl (2011), Lütkepohl and Xu (2012), Proietti and Lütkepohl (2013) and Mayr and Ulbricht (2015) demonstrate that out-of-sample forecasts based on models applied to BC transformed macroeconomic series can be more accurate than those based on using the *original* (non-transformed) series (cf. Nelson and Granger, 1979). Inspired by these results we consider whether BC transformations are useful within the context of forecasting future realised variances.

The use of transformations in the context of realised variance is not new. Indeed, the application of models to log transformed realised variance is common practise; see, e.g., Andersen et al. (2003), Andersen et al. (2007), Corsi (2009), Hansen et al. (2012) and Koopman and Scharth (2013).[1] More recently, BC transformations have been considered in this context; see, e.g., Gonçalves and Meddahi (2011), Weigand (2014), Zheng and Song (2014) and Nugroho and Morimoto (2016).[2] We add to this literature by examining the relative out-of-sample performance of a variety of contemporary models applied to various BC transformed (and original) realised variance measures. In doing this we consider both previously considered transformation parameters (the square root and log transformations), those not previously widely used (the quartic root transformation), and those based on the nature of the data used (that is, an estimated transformation parameter).

The studies conducted by Weigand (2014) and Zheng and Song (2014) come closest to the current paper in that both consider the out-of-sample costs/benefits of applying BC transformations within the context of volatility models. In the former study, Weigand (2014) proposes two BC transformed models of multivariate realised variance (viz. the 'matrix Box-Cox model of realized covariances' and the 'Box-Cox dynamic correlation model'). In the latter study, the framework

[1] The use of log transformed realised variance is based on previous findings that show that these data are Gaussian distributed; see, e.g., Andersen et al. (2001a, 2001b).

[2] Transformations are not always applied. For instance, Bollerslev et al. (2016) augment the popular long-memory heterogenous autoregressive (HAR) model of Corsi (2009), but decide not to apply the log transformation as in the Corsi-proposed HAR model.

builds on the stochastic volatility model proposed by Koopman and Scharth (2013) such that BC transformed realised variance is a linear function of unobserved underlying volatility. Both of these studies indicate that BC transformations (close to the log transformation) are beneficial in this context. We complement and build on these studies in three ways. First, we consider a wide range of univariate realised volatility models, all of which are popular and/or have only recently been proposed. Second, we conduct hypothesis tests that not only examine the relative performance of BC transformed models versus the original model, but also the relative performance of models with different transformations (for instance, log versus quartic root transformations). Third, we analyse whether relative performance is uniform over different realised variance measures, both within a particular market, but also across different market indices. Furthermore, the likely determinants of this variation are investigated.

The question of whether to transform the original realised variance measure will depend on the loss function used to assess forecasting performance. For instance, if one uses the mean square log (MS-log) error loss function (given by the mean of the squared difference between the log forecast and the log realised value), then modelling the log transformed series will deliver the best results as the model parameters are optimised with respect to the same loss function used to assess performance.[3] Thus, to avoid favouring a particular BC transformation in this way, we follow the extant literature and consider the mean square (MS) and quasi-likelihood (QLIK) error loss functions applied to the original realised variance measure. These belong to the Bregman loss function family; see Banerjee et al., (2005), Gneiting (2011) and Patton (2015) for further details.[4] Under these loss functions the optimal forecast is obtained by minimising any Bregman loss function when using the original data.

It is, however, quite possible that the models themselves may not be 'suited' to the original (possibly highly non-Gaussian) data. This leads to the possibility that models applied to transformed data may be superior because they more closely match the true data generating process. For instance, there is considerable evidence that documents increased (long memory) persistence

[3]This implicitly assumes that the parameters are estimated by minimising the sum of squared errors.

[4]The Bregman loss function family possess the quality that for correctly specified models with nested information sets, using the MS loss function to rank forecast quality leads to consistent ranking over all members of the Bregman loss function family (Patton, 2015).

(and hence predictability) under the log transformation assumption; see Ding et al. (1993), Ding and Granger (1996) and Proietti (2016). Consequently, although the parameters in the BC transformed model are not optimised with respect to the loss function used to assess performance, their forecasts may be superior because of the suitability of the model to the BC transformed data. It is this trade-off (parameter optimisation versus model suitability) that we are examining in the current paper by considering whether or not to transform realised variance data.

Using a comprehensive set of realised variance measures, we examine whether the use of BC transformations has value to forecasters. The results indicate that such transformations can improve forecasts of future realised variance across a range of models and under both the MS and QLIK loss functions. Moreover, quality differences between forecasts based on modeling the original and transformed data can be significant even when one controls for data mining by using the reality check statistical tests proposed by White (2000). Of the BC transformations that we consider it is the quartic root (and not the log) transformation that delivers the best results. Finally, we demonstrate that the benefits of BC transformation are not evenly spread over the realised variance measures. Indeed, for some measures no benefits are found – a result that we demonstrate is driven by the degree of skewness in the original realised variance measure.

The rest of the paper is organised as follows. The next section contains a description of the methodologies employed and is followed by the empirical results. The final section concludes.

## 2. Methodologies

This section contains the models and methods used to constructed forecasts, and the means by which the relative quality of the forecasts is assessed.

### *2.1. Forecast construction: The problem*

Let $x_t$ be the original data that we wish to forecast, in our case realised variance, $x_t > 0$ and $t = 1, 2, \ldots, T$. As $x_t$ is likely to be highly non-Gaussian we model the BC transformed data given by

$$y_t = f(x_t; \lambda) = \begin{cases} \frac{x_t^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \\ \ln x_t, & \lambda = 0. \end{cases} \tag{1}$$

It follows that $x_t = g(y_t;\lambda) = f^{-1}(y_t;\lambda)$.[5] Suppose the forecaster models $y_t$ and obtains $h$-step ahead forecasts given by the conditional mean of $y_{t+h}$, that is, $\mathrm{E}[y_{t+h}|\mathcal{F}_t]$, where $\mathcal{F}_t$ is the forecaster's information set. Moreover, suppose that we require the conditional mean of $x_{t+h}$, that is, $\mathrm{E}[x_{t+h}|\mathcal{F}_t]$, or equivalently $\mathrm{E}[g(y_{t+h};\lambda)|\mathcal{F}_t]$.[6]

*2.2. Forecast construction: The solution(s)*

One obvious solution would be to take $g(\mathrm{E}[y_{t+h}|\mathcal{F}_t];\lambda)$, henceforth referred to as the *naive adjustment* forecast. However, Jensen's inequality tells us that for convex functions (like $g$) $g(\mathrm{E}[y_{t+h}|\mathcal{F}_t];\lambda) \le \mathrm{E}[g(y_{t+h};\lambda)|\mathcal{F}_t]$. Therefore an expression for $\mathrm{E}[g(y_{t+h};\lambda)|\mathcal{F}_t]$ is required. To this end, taking conditional expectations of a Taylor series expansion about the conditional mean of $y_{t+h}$ (denoted $\mu_{t+h|t}$) gives

$$\mathrm{E}[g(y_{t+h};\lambda)|\mathcal{F}_t] = g(\mu_{t+h|t};\lambda)\left(1+\sum_{k=1}^{\infty} g_k(\mu_{t+h|t};\lambda)\mu_{k,t+h|t}\right), \tag{2}$$

where $\mu_{k,t+h|t}$ is the $k$th conditional moment of $y_{t+h}$ about its conditional mean, and

$$g_k(\mu_{t+h|t};\lambda) = \frac{1-\lambda(k-1)}{k(1+\lambda\mu_{t+h|t})} g_{k-1}(\mu_{t+h|t};\lambda),$$

with $g_0 = 1$. This is henceforth referred to as the *full adjustment* forecast. Simplifications of the expression in (2) are possible. One could assume that $y_t$ has a Gaussian distribution, or one could ignore all moments except the mean and variance.[7] Adjustments based on these assumptions are provided in Table 1, and deliver forecasts henceforth referred to as the *Gaussian adjustment* and *second-order adjustment* forecasts, respectively.[8]

Insert Table 1 here

[5] Note that $y_t$ represents the original realised variance measure when $\lambda = 1$.

[6] Under the Bregman loss function assumption, the conditional mean is the optimal forecast (Banerjee et al., 2005, Gneiting, 2011, and Patton, 2015).

[7] The Gaussian assumption could be problematic. This is because the transformed series has limited support: specifically, $y_t > -1/\lambda$ for $\lambda > 0$ and $y_t < -1/\lambda$ for $\lambda < 0$. Consequently, unless $\lambda$ equals zero, the BC transformed variable cannot technically be Gaussian (Weigand, 2014).

[8] Subsets of the results in Table 1 have been derived previously. Granger and Newbold (1976) derive the solution associated with the log transformation under the Gaussian distribution assumption via Hermite polynomial expansions, Pankratz and Dudley (1987) obtain the solution when $\lambda = 1/N$, where $N$ is a positive integer, and Proietti and Riani (2009) derive a more general result associated with all $\lambda$ values under the Gaussian distribution assumption.

Without the Gaussianity assumption, implementation of the above formulae require further augmentation in order for them to become practical to use. First, in the subsequent empirical application the full adjustment method employs (2) with a truncation such that only the first ten conditional moments are considered. Extending beyond this point has no effect on the accuracy of forecasts. Second, higher conditional moments (that is, the second conditional moment and higher) are estimated using their unconditional sample counterparts.

### *2.3. Models*

A number of realised variance (and transformations thereof) models are available. We consider a range of models that have recently been proposed in the extant literature. For each model we adopt the lag structure most often advocated or used in previous studies. The following set of models is representative rather than exhaustive.

#### *2.3.1. The HAR model*

The first model considered is the popular (and successful) heterogeneous autoregressive (HAR) model proposed by Corsi (2009). This model provides a parsimonious representation of realised variance (and transformations thereof) that attempts to capture the long memory observed in previous studies; see, e.g., Anderson et al. (2001a, 2001b, 2003). The dynamics of demeaned transformed (daily frequency) realised variance based on the HAR model take the following form:

$$y_t = \gamma_1 y_{t-1} + \gamma_2 \sum_{i=1}^{5} y_{t-i} + \gamma_3 \sum_{j=1}^{22} y_{t-j} + \epsilon_t, \tag{3}$$

where $\epsilon_t$ is a suitably defined (zero-mean) error term. This can be written as a restricted infinite-order autoregressive (AR($\infty$)) model such that conditional expectations are given by

$$\mathrm{E}[y_t|\mathcal{F}_{t-1}] = \sum_{i=1}^{\infty} \pi_i y_{t-i}, \tag{4}$$

where

$$\pi_i = \begin{cases} \gamma_1 + \gamma_2 + \gamma_3, & \text{for } i = 1, \\ \gamma_2 + \gamma_3, & \text{for } i = 2, \dots, 5, \\ \gamma_3 & \text{for } i = 6, \dots, 22, \end{cases}$$

and zero otherwise. It follows from the chain-rule of forecasting that $h$-step ahead forecasts are given by

$$\mathrm{E}[y_{t+h}|\mathcal{F}_t] = \sum_{i=1}^{\infty} \pi_i \mathrm{E}[y_{t+h-i}|\mathcal{F}_t], \tag{5}$$

where $\mathrm{E}[y_{t+j}|\mathcal{F}_t] = y_{t+j}$ for $j \leq 0$, and $y_s = 0$ for $s \leq 0$.[9]

### *2.3.2. The FIMA model*

The next model attempts a more direct representation of the long memory characteristics associated with realised variance (and transformations thereof). In particular, we consider the fractionally integrated moving average (FIMA) model recently considered by Proietti (2016). The model represents an extension of the popular IMA model (that is, the exponential smoothing predictor) often advocated in practice to the fractional case. It is given by

$$(1-L)^d y_t = (1-\theta L)\epsilon_t, \tag{6}$$

where $L$ is the lag operator, $d$ is the (fractional) order of integration, and $\theta$ is the MA coefficient.[10] Rearranging (6) and taking conditional expectations we obtain

$$\mathrm{E}[y_t|\mathcal{F}_{t-1}] = \sum_{i=1}^{\infty} \pi_i y_{t-i}, \tag{7}$$

where $\pi_i$ is the $i$th coefficient in the polynomial:

$$\pi(L) = (1-\theta L)^{-1}(1-L)^d = 1 + \sum_{i=1}^{\infty} \pi_i L^i. \tag{8}$$

It follows that the chain-rule of forecasting can again be applied to yield $h$-step ahead forecasts via an expression analogous to (5).

### *2.3.3. The HEAVY model*

To incorporate improved measures of realised variance (and transformations thereof) into a conditional GARCH-type specification, the first-order high frequency based volatility (HEAVY)

[9] Audrino and Knaus (2016) demonstrate that the restrictive specification of the HAR model is not inferior to an unrestricted AR model in which the specification is determined via the *lasso* selection criterion.

[10] Proietti (2006) provides empirical support for this model over the HAR model.

model proposed by Shephard and Sheppard (2010) takes the form

$$\mathrm{E}[y_t|\mathcal{F}_{t-1}] = \alpha y_{t-1} + \beta \mathrm{E}[y_{t-1}|\mathcal{F}_{t-2}], \tag{9a}$$

$$\mathrm{E}[z_t|\mathcal{F}_{t-1}] = \alpha_R y_{t-1} + \beta_R \mathrm{E}[z_{t-1}|\mathcal{F}_{t-2}]. \tag{9b}$$

Here $z_t$ represents an alternative BC transformed measure of realised variance (given by the BC transformed squared daily return). To differentiate between the two equations, Shephard and Sheppard (2010) refer to the first equation as the HEAVY-RM model and the second as the HEAVY-R model.[11] It is the former that has our focus.

The specification can be written more compactly as

$$\mathrm{E}[\boldsymbol{y}_t|\mathcal{F}_{t-1}] = \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{B}\mathrm{E}[\boldsymbol{y}_{t-1}|\mathcal{F}_{t-2}], \tag{10}$$

where

$$\boldsymbol{y}_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} \alpha & 0 \\ \alpha_R & 0 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \beta & 0 \\ 0 & \beta_R \end{bmatrix}.$$

This equation can also be written in VARMA(1,1) form,

$$\mathrm{E}[\boldsymbol{y}_t|\mathcal{F}_{t-1}] = (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{y}_{t-1} - \boldsymbol{B}\boldsymbol{\epsilon}_{t-1}, \tag{11}$$

where $\boldsymbol{\epsilon}_t = \boldsymbol{y}_t - \mathrm{E}[\boldsymbol{y}_t|\mathcal{F}_{t-1}]$, or as a restricted VAR($\infty$) process such that

$$\mathrm{E}[\boldsymbol{y}_t|\mathcal{F}_{t-1}] = \sum_{i=1}^{\infty} \boldsymbol{\pi}_i \boldsymbol{y}_{t-i}, \tag{12}$$

where $\boldsymbol{\pi}_i = \boldsymbol{A}\boldsymbol{B}^{i-1}$. It also follows that $h$-step ahead forecasts are given by

$$\mathrm{E}[\boldsymbol{y}_{t+h}|\mathcal{F}_t] = \sum_{i=1}^{\infty} \boldsymbol{\pi}_i \mathrm{E}[\boldsymbol{y}_{t+h-i}|\mathcal{F}_t], \tag{13}$$

where $\mathrm{E}[\boldsymbol{y}_{t+j}|\mathcal{F}_t] = \boldsymbol{y}_{t+j}$ for $j \leq 0$, and $\boldsymbol{y}_s = 0$ for $s \leq 0$. Our interest is in the first element of this vector-valued conditional expectation (that is, $\mathrm{E}[y_{t+h}|\mathcal{F}_t]$).

[11]The suffixes RM and R refer to the realised measures and return (squared) equations, respectively.

### 2.3.4. *The RealGARCH model*

As an alternative to the above model, a variant of the RealGARCH model (ignoring leverage effects) proposed by Hansen et al. (2012) can written as

$$\mathrm{E}[y_t|\mathcal{F}_{t-1}] = \delta \mathrm{E}[z_t|\mathcal{F}_{t-1}], \tag{14a}$$

$$\mathrm{E}[z_t|\mathcal{F}_{t-1}] = \alpha_R y_{t-1} + \beta_R \mathrm{E}[z_{t-1}|\mathcal{F}_{t-2}]. \tag{14b}$$

Substituting (14b) into (14a) and rearranging we obtain

$$\mathrm{E}[y_t|\mathcal{F}_{t-1}] = \delta\alpha_R y_{t-1} + \beta_R \mathrm{E}[y_{t-1}|\mathcal{F}_{t-2}], \tag{15a}$$

$$\mathrm{E}[z_t|\mathcal{F}_{t-1}] = \alpha_R y_{t-1} + \beta_R \mathrm{E}[z_{t-1}|\mathcal{F}_{t-2}]. \tag{15b}$$

This can also be written in the matrix form given by (10), but now

$$\boldsymbol{y}_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} \delta\alpha_R & 0 \\ \alpha_R & 0 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \beta_R & 0 \\ 0 & \beta_R \end{bmatrix}.$$

It necessarily follows that this can be expressed as the VARMA(1,1) process in (11) or the restricted VAR($\infty$) process as in (12), ultimately leading to an expression for the $h$-step ahead forecasts.[12] Our focus is again on the first element of the vector-valued conditional expectation.

### 2.4. *Performance assessment*

The performance of the various forecasting methods is assessed using the following homogenous Bregman loss function family proposed in the context of realised variance forecasting by Patton

[12] Comparing these versions of the HEAVY and RealGARCH models we see that the latter is actually a restricted version of the former (with the restrictions involving $\beta$ and $\beta_R$).

(2011):

$$L(x_{t+h}, \widehat{x}_{t+h|t}; b) = \begin{cases} \frac{1}{(b+1)(b+2)}(x_{t+h}^{b+2} - \widehat{x}_{t+h|t}^{b+2}) - \frac{1}{b+1}\widehat{x}_{t+h|t}^{b+1}(x_{t+h} - \widehat{x}_{t+h|t}), & \text{for } b \notin \{-1, -2\}, \\ \widehat{x}_{t+h|t} - x_{t+h} + x_{t+h} \ln\left(\frac{x_{t+h}}{\widehat{x}_{t+h|t}}\right), & \text{for } b = -1, \\ \frac{x_{t+h}}{\widehat{x}_{t+h|t}} - \ln\left(\frac{x_{t+h}}{\widehat{x}_{t+h|t}}\right) - 1, & \text{for } b = -2, \end{cases} \tag{16}$$

where $\widehat{x}_{t+h|t}$ denotes the $h$-step ahead forecast of the original data (that is, realised variance). Here $b = 0$ corresponds to MS loss, and $b = -2$ corresponds to QLIK loss. For instance, in the former case we have

$$L(x_{t+h}, \widehat{x}_{t+h|t}; 0) = \frac{1}{2}(x_{t+h} - \widehat{x}_{t+h|t})^2, \tag{17}$$

which we recognise as MS loss ignoring the multiplicative constant. Under the Bregman loss function family, Patton (2015) demonstrates that the performance rank of a forecasting method can vary over $b$ in the presence of misspecified models, parameter estimation error, or non-nested information sets. Thus we consider performance under both the MS and QLIK loss functions.

The null hypothesis is that models based on transformed data have no superior predictive ability over those based on a competing transformation of the data (including the original data). The alternative hypothesis is that the former do have superior ability. We use differences in the means of the above loss function values to test this null. Formally, we use the reality check approach to the null given by

$$H_0: \max_{k=1,\ldots,K} \mathrm{E}[L_{0,t+h} - L_{k,t+h}] \leq 0, \tag{18a}$$

against the alternative

$$H_1: \max_{k=1,\ldots,K} \mathrm{E}[L_{0,t+h} - L_{k,t+h}] > 0, \tag{18b}$$

where $L_{0,t+h}$ is the forecast loss associated with the benchmark model, and $L_{k,t+h}$ is the forecast loss associated with the $k$th competitor model.

In our case we examine the predictive performance of the set of models based on BC transformed data against a benchmark model based on a competing BC transformation of the data (including the original data). The null hypothesis is that no model based on transformed data outperforms the benchmark model (for a given loss function), against the alternative that at least one model based on transformed data outperforms the benchmark model. To avoid data snooping bias we use the block bootstrap procedure proposed by White (2000) to test the above hypothesis.[13]

## 3. Results

This section contains the empirical results associated with the relative performance of models based on the original and transformed realised variance data.

### *3.1. Data*

We consider ten realised variance measures associated with the S&P 500 index. The first four measures (denoted RV1, RV2, RV3 and RV4) are based on the following estimators: two standard realised variance estimators based on 5 and 10-minute frequency returns (Andersen et al., 2001a, and Barndorff-Nielsen and Shephard, 2002); the jump-robust bipower variation estimator based on 5-minute frequency returns (Barndorff-Nielsen and Shephard, 2004); and the downside risk semivariance estimator based on 5-minute frequency returns (Barndorff-Nielsen et al., 2010). In addition, four multiscale versions of these four estimators are considered (denoted RV5, RV6, RV7 and RV8) in which a 1-minute subsample is used (Zhang et al., 2005). Finally, we consider the microstructure noise-robust realised kernel estimator (Barndorff-Nielsen et al., 2008), and the jump-robust median truncated realised variance estimator (Andersen et al., 2012). These are denoted RV9 and RV10, respectively. Loosely speaking, the RV1 to RV4 measures are more basic in construction than the more complex RV5 to RV10 measures. All series were collected from the Oxford-Man Institute of Quantitative Finance Realized Library (`http://realized.oxford-man.ox.ac.uk/data`). The data span the period from January 1, 2000, to December 31, 2015.

### *3.2. Estimation details*

In addition to estimating the models using the original data ($\lambda = 1$), we consider four different BC transformations. Specifically, the models are estimated using square root transformed data ($\lambda = 1/2$), quartic root transformed data ($\lambda = 1/4$), and log transformed data ($\lambda = 0$). In

[13] We follow White (2000) and make use of the stationary bootstrap proposed by Politis and Romano (1994) in which the bootstrap data are composed of sub-sample blocks with lengths drawn from a geometric distribution with mean equal to a pre-selected value. We assume that 1000 replications of the data are drawn with the mean block length set equal to five observations.

addition, we consider a BC transformation in which the transformation parameter is estimated based on the data used (henceforth $\lambda$-estimated transformed data). We adopt the nonparametric procedure proposed by Proietti and Lütkepohl (2013) in which the Hannan and Nicholls (1977) estimator of the 1-step ahead prediction error variance (p.e.v.) is minimised with respect to $\lambda$.[14] Optimisation of $\lambda$ over this space is achieved using the Constrained Optimisation (CO) package in GAUSS v.11, where solutions are obtained using the Newton-Raphson and cubic/quadratic step length methods.

The models are estimated using the original data and each of the above BC transformations of the data. Estimation is achieved by minimising the sum of squared 1-step ahead prediction errors. For the HAR model ordinary least squares (OLS) is used. By contrast, the FIMA, HEAVY and RealGARCH parameters are estimated using the Newton-Raphson and cubic/quadratic step length methods, again implemented via the CO package in GAUSS v.11. As the HEAVY and RealGARCH models make use of squared returns, applying the log transformation could be problematic as they could equal zero. To overcome this issue we adopt the following procedure: instances of zero squared returns (lagged) are replaced by the realised variance measure (lagged) being modelled.[15] The log transformation is then taken as normal, and estimation is possible.

For the out-of-sample analysis two updating schemes are adopted, with the initial estimation window from January 1, 2000 to December 31, 2005, common to both. The increasing window (denoted Incr-W) scheme adds one observation to the end of the estimation window until the December 31, 2015 observation is reached, while maintaining the same start point. By contrast, the rolling window (denoted Roll-W) scheme also adds one observation to the end of the sample until the December 31, 2015 observation is reached, but removes one observation from the start of the sample.[16] In both cases, out-of-sample 1 to 5-step ahead forecasts of realised variance (that is, $\widehat{x}_{t+h|t}$) are generated at each point (the formulae in Table 1 are used to convert the transformed forecasts to realised variance forecasts).[17] The realised variance forecasts based on models of the

[14]We follow Proietti and Lütkepohl (2013) and use three consecutive raw periodogram ordinates in the construction of the p.e.v.

[15]Alternative approaches are possible. For instance, one could replace zero squared returns with a small positive constant. However, there is unlikely to be any noticeable difference in the results in this particular application as occurrences of zero squared returns are extremely rare. In particular, over the full sample period only one instance occurred within the 4,035 observations.

[16]We follow Corsi (2009) and adopt a 1000-day rolling window.

[17]A two-step procedure is adopted when using $\lambda$-estimated transformed data. First, within each estimation window,

original data are then compared to the realised variance forecasts based on transformed data.

*3.3. Preliminary analysis*

It is interesting to consider the relations between the BC transformation parameter and various aspects of the dynamics of volatility. To this end, Figure 1 contains plots of the p.e.v., the fractional difference parameter ($d$) from the FIMA model, series persistence (given by the sum of the first 50 AR parameters implied by the FIMA model), and the fit associated with the FIMA model (given the $R^2$ statistic), against $\lambda$. These plots are provided for each realised variance measure and for $\lambda$ values between minus one and plus one. The full sample of data is used to estimate the p.e.v. and the parameters of the FIMA model.

Insert Figure 1 here

The plots in Figure 1 reveal a number of interesting characteristics of the data. First, the p.e.v. is minimised around zero – supporting the widespread use of the log transformation. By contrast, the long memory, persistence and model fit values are maximised for $\lambda$ values around 1/4. Thus quartic root transformation appears to deliver data that is most predictable. This does not imply, however, that this transformation will deliver the best forecasts of the original data. There are two reasons for this. First, the plots are based on in-sample estimation, whereas our interest is on out-of-sample forecast quality. Second, we are assessing performance using MS and QLIK loss. The former loss is closely related to the p.e.v., which the quartic root does not appear to minimise (recall the p.e.v. is minimised around $\lambda = 0$). This does, however, leave open the possibility that the quartic root transformation could have value when considering out-of-sample forecasting performance, and/or when using a loss function other than MS loss (that is, QLIK loss). These possibilities will be investigated in the subsequent analysis.

To provide further insight into how model performance varies over different $\lambda$ values, Table 2 contains the estimated parameter and model fit values associated with the FIMA model. All results in this table are based on the full sample of data. Panel A contains the skewness and estimated $\lambda$ values for each realised variance measure, while the other panels contain the FIMA parameter

$\lambda$ is estimated. Second, the parameters associated with each model applied to these $\lambda$-estimated transformed data are then estimated within the same window.

values when original and transformed realised variance data are used. In particular, we consider the original data (panel B), square root transformed data (panel C), quartic root transformed data (panel D), log transformed data (panel E), and $\lambda$-estimated transformed data (panel F). In addition, the $R^2$ statistics are provided. Two sets of these statistics are provided. The first correspond to those observed in the model. The second set are calculated by first transforming the fitted values into the original data form and then calculating the $R^2$ values based on these original fitted values. In doing this we are able to compare consistent measures of fit over the different data used.

Insert Table 2 here

There is variation in the skewness levels associated with each realised variance measure. Perhaps most notable is the increase in skewness as one moves from the basic realised variance measures (RV1 to RV4) to the more sophisticated measures (RV5 to RV10). This suggests that the (high) skewness in realised variance is to some extent obscured by the noise inherent in the basic measures. Somewhat surprisingly this variation does not appear to be as dramatic when the estimated $\lambda$ values are considered. Here most values are close to $-0.03$. This compares to the $\lambda$ values of around $-0.1$ and $-0.05$ estimated by Weigand (2014) and Zheng and Song (2014), respectively.

The results in Table 2 also indicate that the FIMA model parameters are large – a result that highlights the persistent nature of the data. Moreover, the model provides a good fit to the data. For instance, using the realised variance measure based on 5-minute frequency returns (RV1), the $R^2$ statistic equals 54.222%. When BC transformed data are used the fit increases dramatically indicating improved suitability to these data. For instance, the corresponding $R^2$ statistic equals 69.507% when log transformed data are used. To enable an appropriate comparison between these fit measures, we transform the fitted log transformed data back to the fitted original data and recompute the $R^2$ statistic. Doing this using the naive adjustment method delivers an $R^2$ statistic of 51.064%. This value rises to 54.654% when the full adjustment method is used.[18] Importantly this value exceeds that observed when the original data is modeled directly. Thus use of the FIMA model with BC transformed data delivers a superior representation of the data. Similar results hold for the other realised variance measures and BC transformations.

[18]Given the superiority of the full adjustment method we focus exclusively on this method in the subsequent analysis.

*3.4. Out-of-sample performance*

The previous analysis demonstrates that BC transformations have virtue in an in-sample estimation setting. However, the true test of the approach is within an out-of-sample context. To this end, the HAR, FIMA, HEAVY and RealGARCH models are estimated using the original and BC transformed realised variance data. The MS loss associated with each model relative to (divided by) that associated with the HAR model (using Incr-W estimation) applied to the original data are provided in Table 3. In addition to daily horizon forecasts, we also consider weekly horizon forecasts based on integrated 1 to 5-step ahead forecasts. Extant results are provided in Table 4.

Insert Tables 3 and 4 here

The results in Table 3 indicate that there is variation in the performance of models applied to the original data, with the HAR model delivering the least accurate forecasts and the FIMA model delivering the most accurate forecasts (entries below unity). For instance, applying the FIMA model to the realised variance measure based on 5-minute frequency returns (RV1) and using Incr-W estimation delivers a relative MS loss of 0.954. Comparing the results associated with the Incr-W and Roll-W estimation methods it is apparent that the former is universally superior. This relative ranking of the models does not vary considerably over the realised variance measures and loss functions; however, absolute performance does appear to vary over this space.[19]

An obvious question is whether it is better to model original or transformed data in order to deliver improved forecasts of the original data. Applying the models to the BC transformed data delivers more accurate forecasts than those based on the original data, with use of all transformations delivering superior performance. For instance, the square root, quartic root, log and $\lambda$-estimated transformation versions of the FIMA model (using Incr-W estimation) deliver relative MS losses

[19] A number of variants of the models are also considered in the empirical section: viz. the (first-order) autoregressive fractionally integrated (ARFI) model, the (first-order) autoregressive fractionally integrated moving average (ARFIMA) model, and second-order versions of the HEAVY and RealGARCH models in which two lagged values of $y_t$ enter into (9a), (9b) and (14b). Applying these models to the realised variance measure based on 5-minute frequency returns (RV1) and using Incr-W estimation with a daily horizon delivers relative MS losses of 0.979 (ARFI), 0.954 (ARFIMA), 1.062 (second-order HEAVY), and 0.971 (second-order RealGARCH). Comparing with the relative efficiencies associated with the FIMA and first-order HEAVY and RealGARCH models presented in Table 2 (that is, 0.954, 0.979, and 0.965, respectively) we note that the variants are not superior to the models for which results are presented. Consequently, we maintain our use of the smaller set of models in the subsequent analysis.

associated with RV1 of 0.879, 0.875, 0.879 and 0.879, respectively.[20] While the relative performance of the models does not seem to vary over the realised variance measures, it is noticeable that there is still considerable variation in the degree of benefit from BC transforming over this space.

Similar results are observed in Table 4: models based on BC transformed data perform better than those based on the original data. However, now the benefits are greater. For instance, the square root, quartic root, log and $\lambda$-estimated transformation versions of the FIMA model (using Incr-W estimation) now deliver relative MS losses associated with RV1 of 0.763, 0.763, 0.776 and 0.776. Thus modeling transformed data delivers a meaningful improvement in forecasting performance. While the extent of this improvement does depend of the realised variance measure used, it is never detrimental to model BC transformed data.

The effects noted in Tables 3 and 4 are large in magnitude and hence appear economically significant. However, we can go further and subject these to a statistical test. The results in Tables 5 and 6 contain the $p$-values associated with the bootstrap reality check tests described in subsection 2.4. The results are based on a comparison of all models applied to BC transformed data (for instance, all models based on log transformed data with Incr-W and Roll-W estimated parameters) with the benchmark model. Three different benchmark models are considered: the best model applied to the original data (henceforth benchmark $\mathcal{A}$), the best model applied to the quartic root transformed data (benchmark $\mathcal{B}$), and the best model applied to the log transformed data (benchmark $\mathcal{C}$). Daily and weekly horizon results under the MS and QLIK loss function assumptions are provided in Tables 5 and 6, respectively.

Insert Tables 5 and 6 here

A number of findings are apparent in Table 5 (daily horizon). The null that no model based on BC transformed data has superior predictive ability to the best model based on the original data (that is, benchmark $\mathcal{A}$) can be rejected at the 10% level in a large number of instances. However, there is variation in the rejection rates over the realised variance measures, over the loss functions, and over the BC transformation parameters. It is also noteworthy that the quartic root

[20]The performance of the $\lambda$-estimated transformation versions of the models is indistinguishable from the performance of the log transformation versions. This is ultimately because the estimated $\lambda$ values are only slightly less than zero. See Table 2 for in-sample estimates. Time-varying estimates based on Incr-W and Roll-W estimation are available on request.

transformation ($\lambda = 1/4$) performs extremely well.[21] Indeed, it is the only transformation that delivers average $p$-values less than 0.05 under *both* the MS and QLIK loss functions.[22]

The results in Table 6 (weekly horizon) show that the quartic root transformation is still useful, with more rejections of the null (at the 10% level) than those associated with the log transformation under both loss functions. Moreover, the log transformation delivers forecasts that are least attractive overall. In particular, the average $p$-values under this transformation equal 0.088 (MS loss) and 0.463 (QLIK loss). This compares to respective average $p$-values of 0.063 and 0.028 for the quartic root transformation. Thus over daily and weekly horizons the quartic root transformation performs relatively well with respect to benchmark $\mathcal{A}$.

It is perhaps rather unsurprising that when one considers the two other benchmarks (based on BC transformed data), the null is more difficult to reject. Indeed, for daily horizons, all average $p$-values are considerably greater than 0.1 under the benchmark $\mathcal{B}$ and $\mathcal{C}$ assumptions. However, for weekly horizons, performance difference emerge. In particular, the quartic root transformation delivers an average $p$-value of 0.003 under the QLIK loss function and benchmark $\mathcal{C}$ assumptions. By contrast, the log transformation delivers an average $p$-value of 1.000 under the QLIK loss function and benchmark $\mathcal{B}$ assumptions. In summary, the quartic root transformation is not inferior to any transformation *and* occasionally dominates the log transformation.

### *3.5. Performance determinants*

This subsection investigates the potential determinants of the superior performance of forecasting models based on BC transformed data documented above.

#### *3.5.1. Performance over time*

It is firstly interesting to consider whether the benefits of using BC transformed data are constant over time. Figure 2 provides plots of relative mean losses against time when the HAR model (using Incr-W estimation) is applied to quartic root transformed realised variance under the MS and QLIK loss functions and for daily and weekly horizons. Time variation in the mean forecast

[21] This result is not at odds with other studies. For instance, Nugroho and Morimoto (2016) find that a $\lambda$ value around 0.1 is optimal in their BC transformed stochastic volatility model.

[22] The higher rejection rates noted under QLIK loss may reflect the higher test power observed under this loss function (Patton and Sheppard, 2009).

losses is achieved by smoothing forecast losses using a Gaussian kernel smoothing estimator with window size equal to 66 observations.

Insert Figure 2 here

Under the MS loss function assumption the plots indicate that major benefits are available around the high volatility episode observed during the 2008 financial crisis. By contrast, under the QLIK loss function assumption the benefits appear more evenly distributed over time with no noticeable benefit observed around the financial crisis. The QLIK loss function penalises under-prediction more than over-prediction, while the MS loss function is symmetric. It follows that in comparison to the forecasts associated with models applied to the original data, the forecasts associated with models applied to transformed data are generally more accurate but tend to under-predict future realised variance.

*3.5.2. Performance and skewness*

The results in subsection 3.4 show that the BC transformation does not work well for all realised variance measures. The motivation for use of the BC transformation was that the underlying data are likely to be non-Gaussian (with high positive skewness), and not compatible with the models applied to these data. It follows that the benefits from such a transformation are likely to increase (decrease) as skewness increases (decreases). To investigate this prediction we consider the relationship between the mean losses associated with each model applied to each transformed realised variance measure (relative to the mean losses associated with each model applied to each original realised variance measure), *and* the unconditional sample skewness associated with each original realised variance measure. The scatter plots in Figure 3 depict this relationship when the squared root, quartic root and log transformed versions of the HAR model (using Incr-W estimation) are used, under the MS and QLIK loss functions and for daily and weekly horizons. In addition, the OLS fitted values associated with these scatter plots are also presented.

Insert Figure 3 here

The plots depict a clear negative relationship - that is, more (less) skewness is associated with lower (higher) relative losses when using transformed data. The plots also show that this relationship is most acute when the log transformation is used. Moreover, the unreported $p$-values associated with the OLS slope coefficients from regressions of relative mean loss on skewness are uniformly close to zero. This negative relationship holds true under the MS and QLIK loss function assumptions, though the slope is less steep under the latter assumption. A similar conclusion can be drawn when weekly horizons are considered.

*3.5.3. Performance and skewness (alternative indices)*

It is possible to go further and investigate whether the relationship between relative performance and skewness is maintained when forecasting the variance associated with indices other than the S&P 500 index. To this end, we consider ten realised variance measures based on the NASDAQ-100, ten on the Nikkei 225, and ten on the FTSE 100 index. These ten measures correspond to the RV1 to RV10 measures defined previously, and are also collected from the Oxford-Man Institute of Quantitative Finance Realized Library.

The scatter plots in Figure 4 depict the relationship between mean forecast losses and skewness when the quartic root transformed version of the HAR model (using Incr-W estimation) is used, under the MS and QLIK loss functions and for daily and weekly horizons. In addition, fitted values associated with the locally weighted scatter plot smoother (LOWESS) are also presented.[23]

Insert Figure 4 here

The results indicate that the BC transformation still leads to improved forecasts of realised variance, as evinced by the majority of relative forecast losses below unity. One can also see that the improvements are not as great as those associated with use of S&P 500 index data. Moreover, the negative relationship between relative forecast losses and skewness is maintained, with the skewness noticeably lower for the alternative indices. Thus the improvements observed when using these data are modest because they are associated with less extreme skewness levels. This result reinforces the conclusion that use of the BC transformation is dependent of the nature of the underlying data.

[23]The LOWESS regression assumes a degree of smoothing equal to 0.67, a locally linear fit, and robust symmetric weights.

## 4. Conclusions

The costs/benefits of using forecasts based on models applied to BC transformed realised variance are examined. The findings are summarized as follows. First, forecasts based on various models applied to BC transformations of realised variance tend to be more accurate than those based on various models applied to realised variance. Second, the benefits can be significant in an economic and statistical sense. Third, the commonly-used log transformation does not appear to deliver the best results in terms of statistical significance. Rather it is the quartic root transformation that exhibits the best quality in this regard. Finally, relative forecast accuracy varies over the realised variance measures, with data skewness driving this cross-sectional variation. Moreover, performance varies over time and appears to be a function of market conditions (primarily volatility levels).

The results have implications for researchers and market practitioners. Care is required when comparing the performance of proposed models in that BC transformations can have a large impact on results. It seems that some form of BC transformation is beneficial in terms of forecast improvement under most circumstances. However, the scale of the improvement will depend on the nature of the data as not all realised variance measures are suited to data transformation. In particular, only measures with high skewness appear ripe for transformation; see Lütkepohl and Xu (2012) for similar conditional findings in the context of macroeconomic forecasting. Under these conditions we would advise that a quartic root transformation be applied.

## References


Andersen, T., Bollerslev, T. & Diebold, F. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics, 89, 707–720.

Andersen, T., Bollerslev, T., Diebold, F. & Ebens, H. (2001a). The distribution of realized stock return volatility. Journal of Financial Economics, 61, 43–76.

Andersen, T., Bollerslev, T., Diebold, F. & Labys, P. (2001b). The distribution of exchange rate volatility. Journal of the American Statistical Association, 96, 42–55.

Andersen, T., Bollerslev, T., Diebold, F. & Labys, P. (2003). Modeling and forecasting realized volatility. Econometrica, 71, 579–625.

Andersen, T., Dobrev, D. & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics, 169, 75-93.

Audrino, F. & Knaus, S. (2016). Lassoing the HAR model: A model selection perspective on realized volatility dynamics. Econometric Reviews 35, 1485–1521.

Banerjee, A., Guo, X. & Wang, H. (2005). On the optimality of conditional expectation as a Bregman predictor. IEEE Transactions on Information Theory, 51, 2664–2669.

Bårdsen, G. & Lütkepohl, H. (2011). Forecasting levels of log variables in vector autoregressions. International Journal of Forecasting, 27, 1108–1115.

Barndorff-Nielsen, O., Hansen, P., Lunde, A. & Shephard, N. (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica, 76, 1481–1536.

Barndorff-Nielsen, O., Kinnebrouck, S. & Shephard, N. (2010). Measuring downside risk: Realised semivariance. In T. Bollerslev, J. Russell, and M. Watson (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle. Oxford University Press.

Barndorff-Nielsen, O. & Shephard, N. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society, Series B, 64, 253–280.

Barndorff-Nielsen, O. & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps, Journal of Financial Econometrics, 2, 1–37.

Bollerslev, T., Patton, A. & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. Journal of Econometrics, 192, 1–18.

Box G. & Cox, D. (1964). An analysis of transformations (with discussion). Journal of the Royal Statistical Society, Series B, 26, 211–246.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7, 174–196.

Ding, Z., Granger, C., & Engle, R. (1993). A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83–106.

Gneiting, T. (2011). Making and evaluating point forecasts, Journal of the American Statistical Association, 106, 746–762.

Gonçalves, S. & Meddahi, N. (2011). Box-Cox transforms for realized volatility. Journal of Econometrics, 160, 129–144.

Granger, C., & Ding, Z. (1996). Varieties of long memory models. Journal of Econometrics 73, 61–78.

Granger, C. & Newbold, P. (1976). Forecasting transformed time series. Journal of the Royal Statistical Society, Series B, 38, 189–203.

Hannan, E., & Nicholls, D. (1977). The estimation of prediction error variance. Journal of the American Statistical Association, 72, 834–840.

Hansen, P., Huang, Z. & Shek, H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. Journal of Applied Econometrics, 27, 877–906.

Koopman, S. & Scharth, M. (2013). The analysis of stochastic volatility in the presence of daily realized measures. Journal of Financial Econometrics, 11, 76–115.

Lütkepohl, H. & Xu, F. (2012). The role of the log transformation in forecasting economic variables. Empirical Economics, 42, 619–638.

Mayr, J. & Ulbricht, D. (2015). Log versus level in VAR forecasting: 42 million empirical answers - Expect the unexpected. Economics Letters, 126, 40–42.

Nelson, H. & Granger, C. (1979). Experience with using the Box-Cox transformation when forecasting economic time series. Journal of Econometrics, 10, 57–69.

Nugroho, D. & Morimoto, T. (2016). Box-Cox realized asymmetric stochastic volatility models with generalized Students t-error distributions. Journal of Applied Statistics, 43, 1906–1927.

Pankratz, A. & Dudley, U. (1987). Forecasts of power-transformed series. Journal of Forecasting, 6, 239–248.

Patton, A. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160, 246–256.

Patton, A. (2015). Evaluating and comparing possibly misspecified forecasts. Duke University Working Paper.

Patton, A. & Sheppard, K. (2009). Optimal combinations of realised volatility estimators. International Journal of Forecasting, 25, 218–238.

Politis, D. & Romano, J. (1994). The stationary bootstrap. Journal of the American Statistical Association, 89, 1303–1313.

Proietti, T. (2016). Component-wise representations of long-memory models and volatility prediction. Journal of Financial Econometrics, 14, 668–692.

Proietti, T. & Lütkepohl, H. (2013). Does the Box-Cox transformation help in forecasting macroeconomic time series? International Journal of Forecasting, 29, 88–99.

Proietti, T. & Riani, M. (2009). Transformations and seasonal adjustment, Journal of Time Series Analysis, 30, 47–69.

Shephard, N. & Sheppard, K. (2010). Realising the future: Forecasting with high frequency based volatility (HEAVY) models, Journal of Applied Econometrics, 25, 197–231.

Tukey, J. (1957). On the comparative anatomy of transformations, Annals of Mathematical Statistics, 28, 602–632.

Weigand, R. (2014). Matrix Box-Cox models for multivariate realized volatility. University of Regensburg Working Papers in Business, Economics and Management Information Systems.

White, H. (2000). A reality check for data snooping. Econometrica, 68, 1097–1126.

Zhang, L., Mykland, P. & Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association, 100, 1394–1411.

Zheng, T. & Song T. (2014). A realized stochastic volatility model with Box-Cox transformation. Journal of Business and Economic Statistics, 32, 593–605.

**Table 1** – Conditional expectations under BC transformations

| Adjustment | Conditional expectation |
|---|---|
| Panel A. General form | |
| Naive | $g(\mu_{t+h\|t};\lambda)$ |
| Second-order approximation | $g(\mu_{t+h\|t};\lambda)\left(1+\frac{\sigma^2_{t+h\|t}(1-\lambda)}{2(1+\lambda\mu_{t+h\|t})^2}\right)$ |
| Gaussian approximation | $g(\mu_{t+h\|t};\lambda)\left(1+\sum_{k=1}^{\infty} g_{2k}(\mu_{t+h\|t};\lambda)(2k-1)!!\sigma^{2k}_{t+h\|t}\right)$ |
| Full | $g(\mu_{t+h\|t};\lambda)\left(1+\sum_{k=1}^{\infty} g_k(\mu_{t+h\|t};\lambda)\mu_{k,t+h\|t}\right)$ |
| Panel B. Specific form: $N$th root ($\lambda = 1/N$), where $N$ is a positive integer | |
| Naive | $(1+\mu_{t+h\|t}/N)^N$ |
| Second-order approximation | $(1+\mu_{t+h\|t}/N)^N\left(1+\frac{(N-1)\sigma^2_{t+h\|t}}{2N(1+\mu_{t+h\|t}/N)^2}\right)$ |
| Gaussian approximation | $(1+\mu_{t+h\|t}/N)^N\left(1+\sum_{k=1}^{\lfloor N/2\rfloor}\frac{(N-1)!(2k-1)!!\sigma^{2k}_{t+h\|t}}{(2k)!(N-2k)!N^{2k-1}(1+\mu_{t+h\|t}/N)^{2k}}\right)$ |
| Full | $(1+\mu_{t+h\|t}/N)^N\left(1+\sum_{k=1}^{N}\frac{(N-1)!\mu_{k,t+h\|t}}{k!(N-k)!N^{k-1}(1+\mu_{t+h\|t}/N)^{k}}\right)$ |
| Panel C. Specific form: Log ($\lambda = 0$) | |
| Naive | $\exp(\mu_{t+h\|t})$ |
| Second-order approximation | $\exp(\mu_{t+h\|t})\left(1+\frac{\sigma^2_{t+h\|t}}{2}\right)$ |
| Gaussian approximation | $\exp(\mu_{t+h\|t})\exp\left(\frac{\sigma^2_{t+h\|t}}{2}\right)$ |
| Full | $\exp(\mu_{t+h\|t})\left(1+\sum_{k=1}^{\infty}\frac{1}{k!}\mu_{k,t+h\|t}\right)$ |

Notation:
$g(\mu_{t+h|t};\lambda) = (1+\lambda\mu_{t+h|t})^{1/\lambda}$
$g_k(\mu_{t+h|t};\lambda) = \frac{1-\lambda(k-1)}{k(1+\lambda\mu_{t+h|t})} g_{k-1}(\mu_{t+h|t};\lambda)$ and $g_0 = 1$
$\mu_{t+h|t}$ is the $h$-step ahead conditional mean
$\mu_{k,t+h|t}$ is the $h$-step ahead $k$th conditional moment about the conditional mean
$\sigma_{t+h|t}$ $(= \sqrt{\mu_{2,t+h|t}})$ is the $h$-step ahead conditional standard deviation
The notation $\lfloor . \rfloor$ and !! represent the floor and double factorial functions, respectively

Notes: This table contains expressions for the $h$-step ahead conditional expectations of the original data ($x_t$) as a function of the $h$-step ahead conditional moments of the BC transformed data ($y_t$).

**Table 2** – In-sample FIMA model parameter estimates

| Charact./Param. | Realised Variance Measure 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: Characteristics of original data | | | | | | | | | | |
| Skewness | 10.890 | 10.359 | 10.800 | 9.895 | 15.309 | 12.362 | 15.091 | 13.777 | 14.633 | 16.348 |
| BC transformation ($\lambda$) | −0.029 | −0.027 | −0.027 | −0.025 | −0.025 | −0.024 | −0.036 | 0.014 | −0.026 | −0.030 |
| Panel B: Model parameters based on original data | | | | | | | | | | |
| Fractional diff. ($d$) | 0.531 | 0.516 | 0.563 | 0.492 | 0.525 | 0.524 | 0.530 | 0.529 | 0.521 | 0.545 |
| MA coefficient ($\theta$) | 0.182 | 0.159 | 0.202 | 0.250 | 0.221 | 0.199 | 0.230 | 0.257 | 0.205 | 0.255 |
| $R^2$ | 54.222 | 53.216 | 57.367 | 42.701 | 50.174 | 52.240 | 49.647 | 48.035 | 50.714 | 51.103 |
| Panel C: Model parameters based on BC transformed ($\lambda = 1/2$) data | | | | | | | | | | |
| $d$ | 0.584 | 0.584 | 0.597 | 0.538 | 0.593 | 0.583 | 0.594 | 0.563 | 0.590 | 0.610 |
| $\theta$ | 0.163 | 0.209 | 0.097 | 0.240 | 0.099 | 0.112 | 0.098 | 0.170 | 0.131 | 0.091 |
| $R^2$ | 69.738 | 66.588 | 74.675 | 57.315 | 74.663 | 73.003 | 74.460 | 67.109 | 72.682 | 77.055 |
| $R^2$ (naive adj.) | 54.460 | 53.077 | 57.346 | 42.603 | 50.564 | 52.657 | 49.903 | 48.355 | 51.346 | 50.745 |
| $R^2$ (full adj.) | 54.656 | 53.329 | 57.471 | 43.044 | 50.678 | 52.801 | 50.016 | 48.574 | 51.487 | 50.826 |
| Panel D: Model parameters based on BC transformed ($\lambda = 1/4$) data | | | | | | | | | | |
| $d$ | 0.598 | 0.599 | 0.596 | 0.543 | 0.602 | 0.594 | 0.598 | 0.562 | 0.604 | 0.615 |
| $\theta$ | 0.193 | 0.243 | 0.087 | 0.250 | 0.091 | 0.112 | 0.074 | 0.174 | 0.151 | 0.055 |
| $R^2$ | 71.168 | 68.038 | 76.648 | 58.931 | 77.469 | 75.706 | 77.724 | 68.051 | 74.866 | 80.095 |
| $R^2$ (naive adj.) | 53.578 | 51.770 | 56.927 | 40.790 | 50.780 | 52.511 | 49.980 | 47.373 | 51.275 | 50.790 |
| $R^2$ (second-order adj.) | 54.407 | 52.809 | 57.498 | 42.505 | 51.249 | 53.101 | 50.444 | 48.390 | 51.871 | 51.128 |
| $R^2$ (Gaussian adj.) | 54.409 | 52.812 | 57.499 | 42.514 | 51.250 | 53.103 | 50.445 | 48.393 | 51.872 | 51.129 |
| $R^2$ (full adj.) | 54.437 | 52.845 | 57.513 | 42.591 | 51.267 | 53.122 | 50.461 | 48.411 | 51.891 | 51.140 |
| Panel E: Model parameters based on BC transformed ($\lambda = 0$) data | | | | | | | | | | |
| $d$ | 0.591 | 0.588 | 0.583 | 0.527 | 0.593 | 0.583 | 0.586 | 0.545 | 0.600 | 0.604 |
| $\theta$ | 0.200 | 0.240 | 0.090 | 0.241 | 0.096 | 0.113 | 0.069 | 0.180 | 0.164 | 0.059 |
| $R^2$ | 69.507 | 66.428 | 75.055 | 56.807 | 76.116 | 74.335 | 76.636 | 64.835 | 73.347 | 78.586 |
| $R^2$ (naive adj.) | 51.064 | 48.840 | 55.047 | 36.715 | 49.839 | 50.941 | 49.090 | 44.033 | 49.662 | 50.397 |
| $R^2$ (second-order adj.) | 54.358 | 52.771 | 57.661 | 42.683 | 51.895 | 53.443 | 51.131 | 48.421 | 52.177 | 52.051 |
| $R^2$ (Gaussian adj.) | 54.521 | 52.987 | 57.777 | 43.193 | 51.978 | 53.553 | 51.215 | 48.717 | 52.290 | 52.112 |
| $R^2$ (full adj.) | 54.654 | 53.101 | 57.882 | 43.387 | 52.078 | 53.679 | 51.319 | 48.726 | 52.398 | 52.182 |
| Panel F: Model parameters based on BC transformed (estimated $\lambda$) data | | | | | | | | | | |
| $d$ | 0.588 | 0.585 | 0.580 | 0.525 | 0.591 | 0.581 | 0.583 | 0.547 | 0.598 | 0.602 |
| $\theta$ | 0.199 | 0.237 | 0.089 | 0.238 | 0.096 | 0.112 | 0.068 | 0.180 | 0.163 | 0.059 |
| $R^2$ | 69.153 | 66.095 | 74.719 | 56.429 | 75.825 | 74.041 | 76.239 | 65.094 | 73.033 | 78.218 |
| $R^2$ (naive adj.) | 50.617 | 48.380 | 54.677 | 36.148 | 49.610 | 50.656 | 48.757 | 44.307 | 49.355 | 50.181 |
| $R^2$ (second-order adj.) | 54.458 | 52.881 | 57.756 | 42.843 | 51.994 | 53.532 | 51.287 | 48.362 | 52.271 | 52.192 |
| $R^2$ (Gaussian adj.) | 54.678 | 53.158 | 57.917 | 43.479 | 52.105 | 53.679 | 51.411 | 48.608 | 52.424 | 52.279 |
| $R^2$ (full adj.) | 54.803 | 53.252 | 58.028 | 43.620 | 52.211 | 53.812 | 51.519 | 48.630 | 52.536 | 52.350 |

Notes: This table contains the estimated skewness and BC transformation ($\lambda$) values (based on minimising the Hannan-Nicholls nonparametric estimator of the 1-step ahead predictive error variance), and the parameter estimates associated with the FIMA model (that is, the fractional difference parameter, $d$, and the MA coefficient, $\theta$) applied to the original and BC transformed realised variance. Two sets of $R^2$ statistics are provided. The first correspond to those observed when the models are estimated. The second set are calculated by first transforming the fitted values into the original data form and then calculating the $R^2$ values based on these fitted values. Four versions of the second set are provided, each one corresponding to a different way of converting (adjusting) the transformed fitted values to their original data form equivalents.

**Table 3** – Out-of-sample forecast losses (daily horizon)

| | | Realised Variance Measure | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Estimation | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Panel A: Forecast losses using original data | | | | | | | | | | | |
| HAR | Incr-W | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FIMA | | 0.954 | 0.965 | 0.934 | 0.982 | 0.899 | 0.945 | 0.884 | 0.924 | 0.920 | 0.795 |
| HEAVY | | 0.979 | 0.990 | 0.968 | 0.988 | 0.943 | 0.975 | 0.932 | 0.962 | 0.954 | 0.891 |
| RealGARCH | | 0.965 | 0.960 | 0.994 | 0.994 | 0.931 | 0.951 | 0.919 | 0.877 | 0.934 | 0.943 |
| HAR | Roll-W | 1.083 | 1.109 | 1.152 | 1.073 | 1.193 | 1.202 | 1.223 | 1.207 | 1.165 | 1.090 |
| FIMA | | 0.990 | 1.025 | 0.985 | 1.010 | 0.955 | 1.011 | 0.937 | 1.010 | 0.983 | 0.808 |
| HEAVY | | 1.059 | 1.086 | 1.092 | 1.038 | 1.096 | 1.127 | 1.085 | 1.104 | 1.095 | 0.955 |
| RealGARCH | | 1.161 | 1.127 | 1.195 | 1.093 | 1.168 | 1.232 | 1.143 | 1.101 | 1.175 | 1.078 |
| Panel B: Forecast losses using BC transformed data ($\lambda = 1/2$) | | | | | | | | | | | |
| HAR | Incr-W | 0.895 | 0.942 | 0.863 | 0.924 | 0.759 | 0.841 | 0.756 | 0.772 | 0.784 | 0.688 |
| FIMA | | 0.879 | 0.930 | 0.846 | 0.928 | 0.742 | 0.833 | 0.737 | 0.769 | 0.769 | 0.662 |
| HEAVY | | 0.886 | 0.941 | 0.850 | 0.922 | 0.744 | 0.837 | 0.737 | 0.771 | 0.771 | 0.663 |
| RealGARCH | | 0.883 | 0.946 | 0.845 | 0.919 | 0.731 | 0.830 | 0.728 | 0.757 | 0.761 | 0.660 |
| HAR | Roll-W | 0.933 | 0.969 | 0.917 | 0.948 | 0.836 | 0.909 | 0.836 | 0.831 | 0.849 | 0.747 |
| FIMA | | 0.901 | 0.942 | 0.882 | 0.943 | 0.798 | 0.882 | 0.796 | 0.815 | 0.813 | 0.705 |
| HEAVY | | 0.913 | 0.957 | 0.894 | 0.942 | 0.807 | 0.892 | 0.803 | 0.821 | 0.824 | 0.711 |
| RealGARCH | | 0.908 | 0.949 | 0.885 | 0.938 | 0.791 | 0.880 | 0.784 | 0.786 | 0.811 | 0.725 |
| Panel C: Forecast losses using BC transformed data ($\lambda = 1/4$) | | | | | | | | | | | |
| HAR | Incr-W | 0.883 | 0.944 | 0.838 | 0.920 | 0.718 | 0.814 | 0.715 | 0.751 | 0.751 | 0.632 |
| FIMA | | 0.875 | 0.939 | 0.831 | 0.926 | 0.709 | 0.813 | 0.704 | 0.755 | 0.744 | 0.617 |
| HEAVY | | 0.878 | 0.944 | 0.829 | 0.918 | 0.708 | 0.813 | 0.701 | 0.751 | 0.745 | 0.615 |
| RealGARCH | | 0.885 | 0.957 | 0.831 | 0.917 | 0.704 | 0.815 | 0.697 | 0.748 | 0.746 | 0.608 |
| HAR | Roll-W | 0.892 | 0.939 | 0.860 | 0.923 | 0.750 | 0.839 | 0.752 | 0.767 | 0.773 | 0.663 |
| FIMA | | 0.875 | 0.928 | 0.840 | 0.923 | 0.730 | 0.826 | 0.728 | 0.762 | 0.756 | 0.637 |
| HEAVY | | 0.881 | 0.937 | 0.843 | 0.922 | 0.732 | 0.830 | 0.729 | 0.763 | 0.759 | 0.639 |
| RealGARCH | | 0.883 | 0.948 | 0.837 | 0.919 | 0.718 | 0.824 | 0.711 | 0.754 | 0.753 | 0.631 |
| Panel D: Forecast losses using BC transformed data ($\lambda = 0$) | | | | | | | | | | | |
| HAR | Incr-W | 0.881 | 0.945 | 0.825 | 0.914 | 0.699 | 0.801 | 0.694 | 0.742 | 0.739 | 0.607 |
| FIMA | | 0.879 | 0.945 | 0.827 | 0.923 | 0.698 | 0.807 | 0.691 | 0.755 | 0.739 | 0.600 |
| HEAVY | | 0.876 | 0.944 | 0.819 | 0.911 | 0.693 | 0.801 | 0.684 | 0.744 | 0.736 | 0.596 |
| RealGARCH | | 0.890 | 0.970 | 0.826 | 0.906 | 0.695 | 0.810 | 0.683 | 0.740 | 0.744 | 0.594 |
| HAR | Roll-W | 0.883 | 0.940 | 0.835 | 0.919 | 0.710 | 0.808 | 0.710 | 0.745 | 0.744 | 0.619 |
| FIMA | | 0.873 | 0.934 | 0.825 | 0.919 | 0.701 | 0.806 | 0.696 | 0.749 | 0.738 | 0.602 |
| HEAVY | | 0.876 | 0.941 | 0.825 | 0.914 | 0.701 | 0.808 | 0.697 | 0.747 | 0.740 | 0.603 |
| RealGARCH | | 0.888 | 0.965 | 0.834 | 0.917 | 0.704 | 0.815 | 0.694 | 0.749 | 0.746 | 0.607 |
| Panel E: Forecast losses using BC transformed data (estimated $\lambda$) | | | | | | | | | | | |
| HAR | Incr-W | 0.882 | 0.946 | 0.824 | 0.915 | 0.697 | 0.799 | 0.692 | 0.744 | 0.739 | 0.608 |
| FIMA | | 0.879 | 0.945 | 0.826 | 0.921 | 0.696 | 0.805 | 0.689 | 0.756 | 0.739 | 0.601 |
| HEAVY | | 0.876 | 0.945 | 0.817 | 0.911 | 0.691 | 0.800 | 0.681 | 0.746 | 0.736 | 0.597 |
| RealGARCH | | 0.892 | 0.976 | 0.824 | 0.903 | 0.694 | 0.809 | 0.680 | 0.741 | 0.744 | 0.594 |
| HAR | Roll-W | 0.894 | 0.945 | 0.849 | 0.927 | 0.712 | 0.816 | 0.715 | 0.749 | 0.747 | 0.624 |
| FIMA | | 0.876 | 0.936 | 0.829 | 0.921 | 0.700 | 0.807 | 0.697 | 0.753 | 0.738 | 0.604 |
| HEAVY | | 0.880 | 0.943 | 0.827 | 0.917 | 0.700 | 0.809 | 0.696 | 0.750 | 0.739 | 0.604 |
| RealGARCH | | 0.897 | 0.972 | 0.848 | 0.919 | 0.708 | 0.824 | 0.700 | 0.751 | 0.750 | 0.613 |

Notes: This table contains the mean forecast losses for each model relative to the HAR model (using Incr-W estimation applied to the original data). The MS loss function is assumed. Entries below (above) unity indicate superior (inferior) relative performance.

**Table 4** – Out-of-sample forecast losses (weekly horizon)

| Model | Estimation | Realised Variance Measure | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Panel A: Forecast losses using the original data | | | | | | | | | | | |
| HAR | Incr-W | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FIMA | | 0.882 | 0.894 | 0.799 | 0.935 | 0.674 | 0.808 | 0.623 | 0.709 | 0.751 | 0.378 |
| HEAVY | | 1.022 | 1.003 | 0.981 | 0.989 | 1.016 | 1.005 | 0.973 | 0.996 | 1.028 | 0.929 |
| RealGARCH | | 1.014 | 0.954 | 1.030 | 1.060 | 1.039 | 0.995 | 0.996 | 0.863 | 1.021 | 1.123 |
| HAR | Roll-W | 1.345 | 1.405 | 1.681 | 1.244 | 2.046 | 1.952 | 2.159 | 2.049 | 1.856 | 1.381 |
| FIMA | | 0.949 | 1.024 | 0.869 | 0.973 | 0.776 | 0.937 | 0.710 | 0.851 | 0.875 | 0.392 |
| HEAVY | | 1.498 | 1.538 | 1.730 | 1.242 | 2.282 | 2.089 | 2.238 | 2.056 | 2.142 | 1.373 |
| RealGARCH | | 1.760 | 1.671 | 2.030 | 1.398 | 2.939 | 2.609 | 2.838 | 2.209 | 2.711 | 1.900 |
| Panel B: Forecast losses using BC transformed data ($\lambda = 1/2$) | | | | | | | | | | | |
| HAR | Incr-W | 0.777 | 0.845 | 0.694 | 0.835 | 0.479 | 0.676 | 0.455 | 0.528 | 0.544 | 0.277 |
| FIMA | | 0.763 | 0.821 | 0.680 | 0.850 | 0.473 | 0.666 | 0.450 | 0.527 | 0.536 | 0.272 |
| HEAVY | | 0.795 | 0.855 | 0.708 | 0.832 | 0.498 | 0.692 | 0.469 | 0.535 | 0.561 | 0.290 |
| RealGARCH | | 0.787 | 0.852 | 0.703 | 0.835 | 0.486 | 0.682 | 0.462 | 0.527 | 0.550 | 0.289 |
| HAR | Roll-W | 0.859 | 0.937 | 0.797 | 0.892 | 0.594 | 0.798 | 0.583 | 0.625 | 0.647 | 0.342 |
| FIMA | | 0.803 | 0.867 | 0.746 | 0.887 | 0.570 | 0.757 | 0.548 | 0.599 | 0.610 | 0.320 |
| HEAVY | | 0.888 | 0.957 | 0.821 | 0.901 | 0.641 | 0.837 | 0.617 | 0.654 | 0.691 | 0.364 |
| RealGARCH | | 0.861 | 0.932 | 0.795 | 0.889 | 0.613 | 0.809 | 0.587 | 0.603 | 0.662 | 0.369 |
| Panel C: Forecast losses using BC transformed data ($\lambda = 1/4$) | | | | | | | | | | | |
| HAR | Incr-W | 0.759 | 0.827 | 0.668 | 0.832 | 0.446 | 0.651 | 0.420 | 0.511 | 0.516 | 0.245 |
| FIMA | | 0.763 | 0.821 | 0.675 | 0.852 | 0.450 | 0.657 | 0.424 | 0.523 | 0.519 | 0.248 |
| HEAVY | | 0.768 | 0.831 | 0.681 | 0.823 | 0.457 | 0.662 | 0.430 | 0.513 | 0.524 | 0.253 |
| RealGARCH | | 0.763 | 0.828 | 0.677 | 0.829 | 0.451 | 0.657 | 0.426 | 0.513 | 0.519 | 0.250 |
| HAR | Roll-W | 0.788 | 0.863 | 0.697 | 0.845 | 0.474 | 0.683 | 0.449 | 0.528 | 0.543 | 0.260 |
| FIMA | | 0.768 | 0.831 | 0.679 | 0.846 | 0.462 | 0.668 | 0.437 | 0.525 | 0.529 | 0.253 |
| HEAVY | | 0.802 | 0.873 | 0.716 | 0.852 | 0.495 | 0.707 | 0.469 | 0.543 | 0.560 | 0.273 |
| RealGARCH | | 0.794 | 0.870 | 0.706 | 0.845 | 0.480 | 0.692 | 0.454 | 0.531 | 0.548 | 0.269 |
| Panel D: Forecast losses using BC transformed data ($\lambda = 0$) | | | | | | | | | | | |
| HAR | Incr-W | 0.757 | 0.819 | 0.666 | 0.828 | 0.440 | 0.648 | 0.413 | 0.513 | 0.511 | 0.242 |
| FIMA | | 0.776 | 0.829 | 0.693 | 0.866 | 0.454 | 0.671 | 0.427 | 0.539 | 0.524 | 0.250 |
| HEAVY | | 0.759 | 0.819 | 0.676 | 0.814 | 0.447 | 0.655 | 0.421 | 0.511 | 0.514 | 0.248 |
| RealGARCH | | 0.758 | 0.824 | 0.672 | 0.829 | 0.443 | 0.653 | 0.417 | 0.517 | 0.512 | 0.245 |
| HAR | Roll-W | 0.779 | 0.850 | 0.683 | 0.847 | 0.452 | 0.667 | 0.425 | 0.522 | 0.526 | 0.245 |
| FIMA | | 0.782 | 0.841 | 0.693 | 0.868 | 0.458 | 0.680 | 0.429 | 0.539 | 0.531 | 0.247 |
| HEAVY | | 0.789 | 0.856 | 0.710 | 0.853 | 0.472 | 0.695 | 0.447 | 0.541 | 0.540 | 0.254 |
| RealGARCH | | 0.794 | 0.864 | 0.702 | 0.875 | 0.469 | 0.690 | 0.439 | 0.545 | 0.539 | 0.256 |
| Panel E: Forecast losses using BC transformed data (estimated $\lambda$) | | | | | | | | | | | |
| HAR | Incr-W | 0.758 | 0.821 | 0.665 | 0.829 | 0.440 | 0.648 | 0.412 | 0.514 | 0.511 | 0.242 |
| FIMA | | 0.776 | 0.829 | 0.694 | 0.863 | 0.454 | 0.672 | 0.427 | 0.539 | 0.524 | 0.250 |
| HEAVY | | 0.758 | 0.818 | 0.674 | 0.811 | 0.446 | 0.654 | 0.419 | 0.512 | 0.514 | 0.248 |
| RealGARCH | | 0.760 | 0.831 | 0.672 | 0.826 | 0.443 | 0.653 | 0.416 | 0.517 | 0.512 | 0.244 |
| HAR | Roll-W | 0.792 | 0.855 | 0.693 | 0.858 | 0.455 | 0.676 | 0.429 | 0.524 | 0.530 | 0.248 |
| FIMA | | 0.786 | 0.843 | 0.698 | 0.871 | 0.461 | 0.685 | 0.431 | 0.538 | 0.533 | 0.249 |
| HEAVY | | 0.791 | 0.857 | 0.713 | 0.855 | 0.473 | 0.698 | 0.449 | 0.541 | 0.540 | 0.255 |
| RealGARCH | | 0.823 | 0.876 | 0.727 | 0.886 | 0.481 | 0.713 | 0.451 | 0.541 | 0.555 | 0.263 |

Notes: This table contains the mean forecast losses for each model relative to the HAR model (using Incr-W estimation applied to the original data). The MS loss function is assumed. Entries below (above) unity indicate superior (inferior) relative performance.

**Table 5** – Reality check results (daily horizon)

| Loss | Benchmark | Realised Variance Measure 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Av. $p$-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: Test $p$-values using BC transformed data ($\lambda = 1/2$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.031 | 0.026 | 0.026 | 0.003 | 0.017 | 0.023 | 0.015 | 0.024 | 0.025 | 0.023 | 0.021 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 0.001 | 1.000 | 0.053 | 1.000 | 0.722 | 0.000 | 0.678 |
| MS | $\mathcal{B}$ | 1.000 | 0.922 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.992 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| MS | $\mathcal{C}$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Panel B: Test $p$-values using BC transformed data ($\lambda = 1/4$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.065 | 0.039 | 0.049 | 0.002 | 0.052 | 0.050 | 0.032 | 0.066 | 0.069 | 0.033 | 0.046 |
| QLIK | | 0.000 | 0.000 | 0.000 | 0.021 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 |
| MS | $\mathcal{C}$ | 1.000 | 0.576 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.958 |
| QLIK | | 0.929 | 1.000 | 0.253 | 1.000 | 0.828 | 0.980 | 0.990 | 0.697 | 0.960 | 0.707 | 0.834 |
| Panel C: Test $p$-values using BC transformed data ($\lambda = 0$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.084 | 0.239 | 0.044 | 0.004 | 0.052 | 0.054 | 0.051 | 0.075 | 0.071 | 0.055 | 0.073 |
| QLIK | | 0.000 | 0.004 | 0.000 | 0.002 | 0.000 | 0.000 | 0.000 | 0.005 | 0.003 | 0.000 | 0.001 |
| MS | $\mathcal{B}$ | 0.972 | 1.000 | 0.482 | 0.171 | 0.100 | 0.193 | 0.162 | 0.231 | 0.552 | 0.203 | 0.407 |
| QLIK | | 1.000 | 0.455 | 1.000 | 0.655 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.911 |
| Panel D: Test $p$-values using BC transformed data (estimated $\lambda$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.132 | 0.419 | 0.041 | 0.003 | 0.053 | 0.060 | 0.056 | 0.080 | 0.083 | 0.059 | 0.099 |
| QLIK | | 0.000 | 0.000 | 0.000 | 0.019 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 |
| MS | $\mathcal{B}$ | 1.000 | 1.000 | 0.962 | 0.939 | 0.978 | 0.990 | 0.959 | 1.000 | 0.991 | 1.000 | 0.982 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.991 | 1.000 | 1.000 | 0.999 |
| MS | $\mathcal{C}$ | 1.000 | 1.000 | 0.433 | 0.190 | 0.102 | 0.302 | 0.119 | 0.445 | 0.716 | 0.151 | 0.446 |
| QLIK | | 1.000 | 0.806 | 1.000 | 0.982 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.979 |

Notes: This table contains the $p$-values associated with reality check tests of the null hypothesis that no model based on BC transformed data outperforms the benchmark model, against the alternative that at least one model based on BC transformed data outperforms the benchmark model. Three benchmark models are considered: benchmark $\mathcal{A}$ is the best model based on original data, benchmark $\mathcal{B}$ is the best model based on BC (quartic root) transformed data ($\lambda = 1/4$), and benchmark $\mathcal{C}$ is the best model based on BC (log) transformed data ($\lambda = 0$). The final column contains the average $p$-value (Av. p-value) across all realised variance measures.

**Table 6** – Reality check results (weekly horizon)

| | | Realised Variance Measure | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Loss | Benchmark | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Av. $p$-value |
| Panel A: Test $p$-values using BC transformed data ($\lambda = 1/2$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.034 | 0.018 | 0.039 | 0.000 | 0.009 | 0.000 | 0.051 | 0.040 | 0.030 | 0.071 | 0.029 |
| QLIK | | 0.112 | 0.979 | 0.404 | 1.000 | 0.000 | 0.010 | 0.000 | 1.000 | 0.000 | 0.000 | 0.351 |
| MS | $\mathcal{B}$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| QLIK | | 0.597 | 1.000 | 0.423 | 0.951 | 0.991 | 1.000 | 1.000 | 0.992 | 1.000 | 0.926 | 0.888 |
| MS | $\mathcal{C}$ | 1.000 | 0.993 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Panel B: Test $p$-values using BC transformed data ($\lambda = 1/4$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.070 | 0.111 | 0.072 | 0.000 | 0.051 | 0.073 | 0.052 | 0.068 | 0.030 | 0.101 | 0.063 |
| QLIK | | 0.000 | 0.000 | 0.057 | 0.013 | 0.000 | 0.000 | 0.000 | 0.211 | 0.000 | 0.000 | 0.028 |
| MS | $\mathcal{C}$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| QLIK | | 0.000 | 0.000 | 0.011 | 0.000 | 0.000 | 0.020 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 |
| Panel C: Test $p$-values using BC transformed data ($\lambda = 0$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.017 | 0.093 | 0.058 | 0.000 | 0.121 | 0.070 | 0.104 | 0.098 | 0.101 | 0.212 | 0.088 |
| QLIK | | 1.000 | 0.534 | 1.000 | 1.000 | 0.020 | 0.028 | 0.022 | 1.000 | 0.000 | 0.020 | 0.463 |
| MS | $\mathcal{B}$ | 0.967 | 0.991 | 0.947 | 0.688 | 0.745 | 0.919 | 0.577 | 1.000 | 0.782 | 0.889 | 0.852 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Panel D: Test $p$-values using BC transformed data (estimated $\lambda$) | | | | | | | | | | | | |
| MS | $\mathcal{A}$ | 0.021 | 0.100 | 0.093 | 0.000 | 0.060 | 0.111 | 0.100 | 0.123 | 0.078 | 0.110 | 0.080 |
| QLIK | | 1.000 | 0.909 | 1.000 | 1.000 | 0.051 | 0.020 | 0.061 | 1.000 | 0.000 | 0.030 | 0.507 |
| MS | $\mathcal{B}$ | 1.000 | 0.986 | 0.994 | 0.990 | 1.000 | 0.988 | 0.991 | 1.000 | 1.000 | 1.000 | 0.995 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.993 | 1.000 | 1.000 | 0.999 |
| MS | $\mathcal{C}$ | 0.990 | 0.980 | 0.970 | 0.778 | 0.788 | 0.990 | 0.485 | 1.000 | 0.869 | 0.859 | 0.871 |
| QLIK | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Notes: This table contains the $p$-values associated with reality check tests of the null hypothesis that no model based on BC transformed data outperforms the benchmark model, against the alternative that at least one model based on BC transformed data outperforms the benchmark model. Three benchmark models are considered: benchmark $\mathcal{A}$ is the best model based on original data, benchmark $\mathcal{B}$ is the best model based on BC (quartic root) transformed data ($\lambda = 1/4$), and benchmark $\mathcal{C}$ is the best model based on BC (log) transformed data ($\lambda = 0$). The final column contains the average $p$-value (Av. p-value) across all realised variance measures.

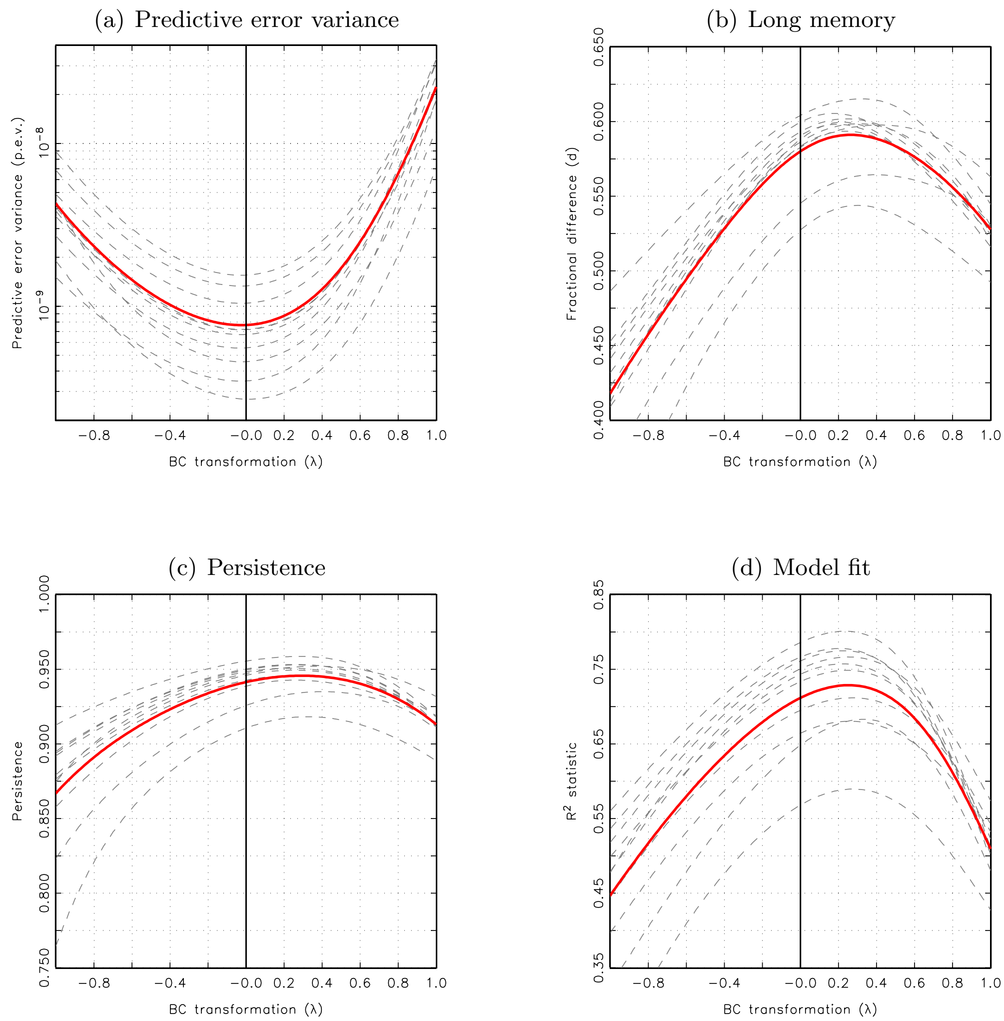


**Figure 1** – Volatility dynamics and the BC transformation parameter ($\lambda$)
This figure contains plots of the estimated 1-step ahead predictive error variance (p.e.v., as given the Hannan-Nicholls non-parametric estimator), the fractional difference parameter ($d$ in the FIMA model), persistence (given by the sum of the first 50 AR coefficients implied by the FIMA model), and model fit ($R^2$ in the FIMA model) against the BC transformation parameter ($\lambda$). Solid lines represent the mean values across all realised variance measures, and the dashed lines are the individual values for each realised variance measure.

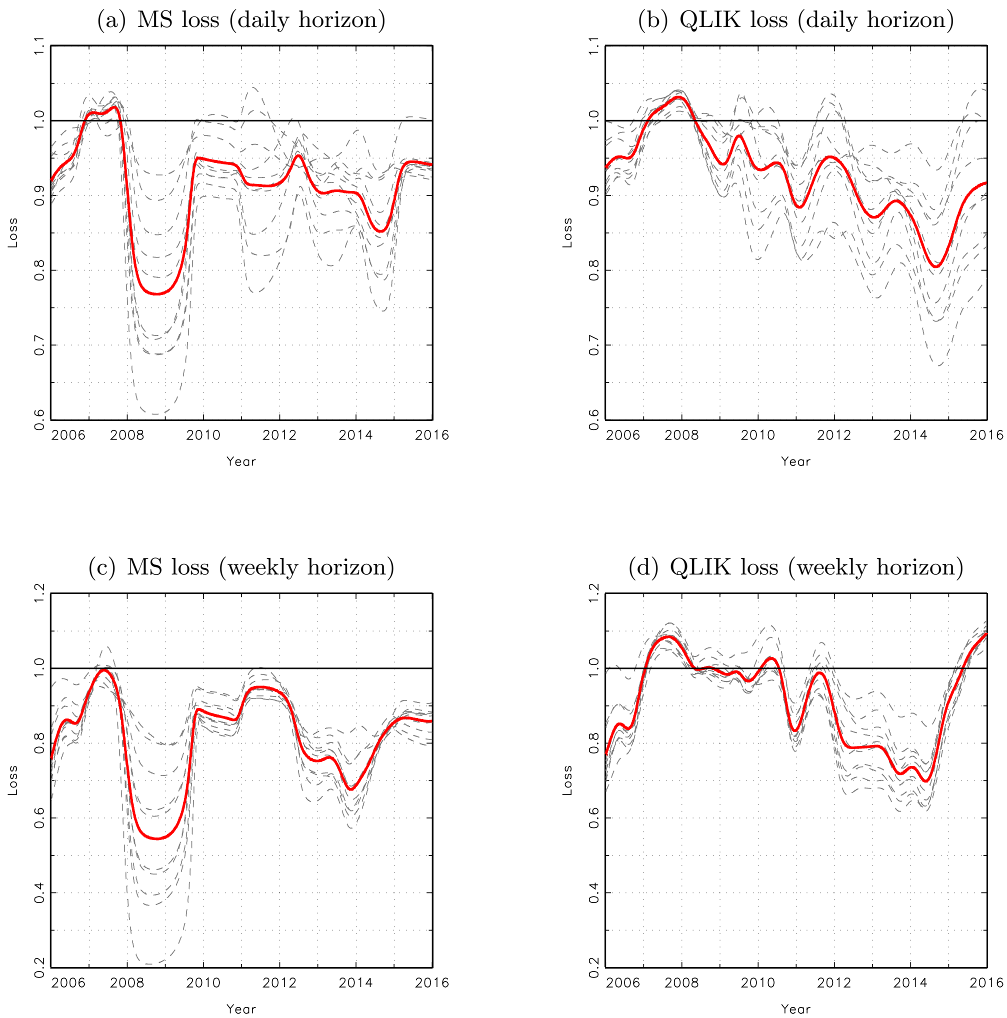


**Figure 2** – Forecast losses over time
This figure contains plots of the time-varying mean forecast losses for quartic root transformed HAR models relative to the HAR model (both using Incr-W estimation). Solid lines represent the relative mean forecast losses across all realised variance measures, and the dashed lines are the individual relative mean forecast losses for each realised variance measure.

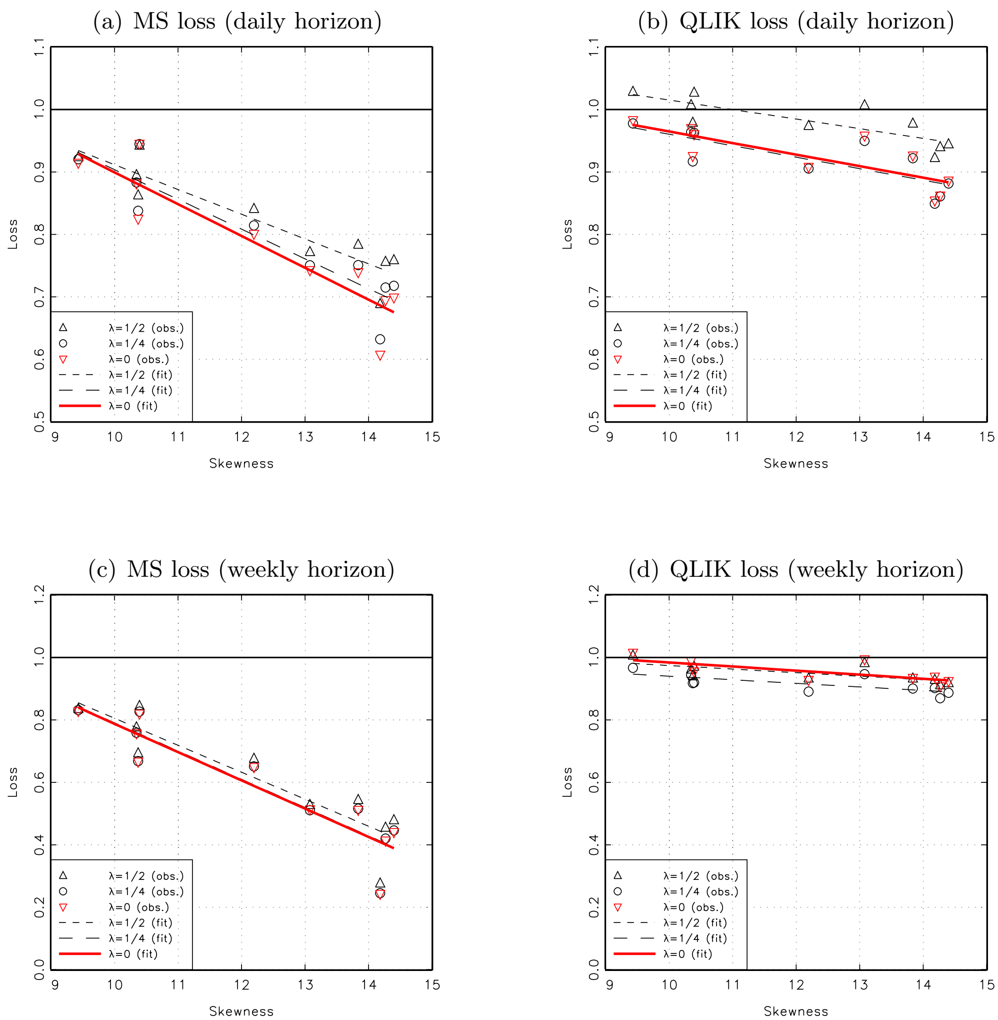


**Figure 3** – Forecast losses and skewness
This figure contains plots of the mean forecast losses for BC transformed HAR models relative to the HAR model (both using Incr-W estimation) against unconditional sample skewness for each realised variance measure. The fit is based on the OLS cross-sectional regression of mean forecast losses on skewness.

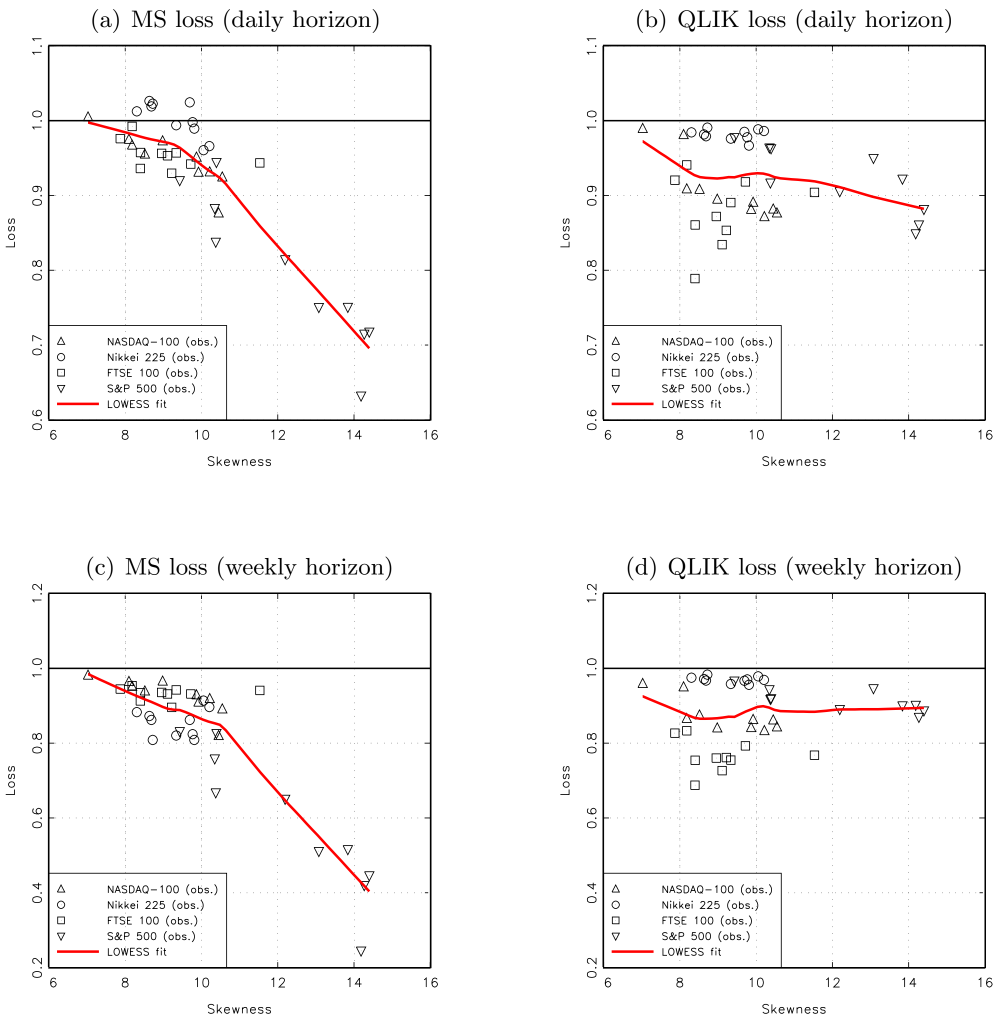


**Figure 4** – Forecast losses and skewness (alternative indices)
This figure contains plots of the mean forecast losses for quartic root transformed HAR models relative to the HAR model (both using Incr-W estimation) against unconditional sample skewness for each realised variance measure. The fit is based on the LOWESS cross-sectional regression of mean forecast losses on skewness.